

22U258

(Pages: 2)

Name:

Reg.No:

SECOND SEMESTER B.Voc. DEGREE EXAMINATION, APRIL 2023

(CBCSS - UG)

(Regular/Supplementary/Improvement)

CC21U SDC2 DS06 - DATA SCIENCE WITH PYTHON

(Information Technology)

(2021 Admission onwards)

Time : 2.5 Hours

Maximum : 80 Marks

Credit : 4

Part A (Short answer questions)

Answer *all* questions. Each question carries 2 marks.

1. What is statistical modeling?
2. Define Exploratory data analysis.
3. How can you select k for k-Means?
4. What are the advantages of K-nearest neighbors' algorithm
5. List the basic principles of data visualization.
6. Create a 3x3 array of normally distributed random values with mean 0 and standard deviation 1.
7. What is the use of np.partition()?
8. What are the rules of broadcasting?
9. What is the use of varify_integrity flag?
10. Define the syntax of series object.
11. What is the use of pandas.eval()?
12. List the pandas str methods.
13. What is Matplotlib in Python?
14. What is the use of alpha keyword?
15. What is the difference between plt.scatter and plt.plot?

(Ceiling: 25 Marks)

Part B (Paragraph questions)

Answer *all* questions. Each question carries 5 marks.

16. Explain basic principles of data visualization.
17. Explain the concatenation and splitting of arrays.

18. Define masking operation.
19. Explain the different aggregate functions in pandas.
20. Explain the merge operation in pandas.
21. Explain Data Indexing in pandas.
22. How to handle null values in pandas?
23. Explain reindexing in pandas.

(Ceiling: 35 Marks)

Part C (Essay questions)

Answer any *two* questions. Each question carries 10 marks.

24. What is data science and explain the data science process?
25. Explain linear regression with suitable example.
26. What is probability? Explain probability distribution in detail.
27. Discuss the different computations performed on NumPy arrays.

(2 × 10 = 20 Marks)
